AD_____

Award Number: DAMD17-01-1-0729

TITLE: Molecular Epidemiology of Ovarian Cancer

PRINCIPAL INVESTIGATOR: David Bowtell, Ph.D.
Adele Green, Ph.D.
Georgia Chenevix-Trench, Ph.D.
Anna DeFazio, Ph.D.
Dorota Gertig, MBBS, Sc.D.
David Purdie, Ph.D.
Penelope Webb, Ph.D.
David Whiteman, Ph.D.

CONTRACTING ORGANIZATION: University of Melbourne
East Melbourne Victoria 3002

REPORT DATE: September 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20060315 076

# REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

| 1. REPORT DATE (DD-MM-YYYY) 01-09-2005 | 2. REPORT TYPE Annual | 3. DATES COVERED (From - To) 15 Aug 04 – 14 Aug 05 |
|---|---|---|

**4. TITLE AND SUBTITLE**
Molecular Epidemiology of Ovarian Cancer

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
DAMD17-01-1-0729

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
David Bowtell, Ph.D., Adele Green, Ph.D., Georgia Chenevix-Trench, Ph.D.,
Anna DeFazio, Ph.D., Dorota Gertig, MBBS, Sc.D., David Purdie, Ph.D.,
Penelope Webb, Ph.D., David Whiteman, Ph.D.

E-Mail: david.bowtell@petermac.org

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

University of Melbourne
East Melbourne Victoria 3002

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

The aim of this Program is to study the association between epidemiologic risk factors, low-risk genes, and histologic and novel molecular subtypes of ovarian cancer. In December 2002, we received final approval from the Human Subject Research Review Board (HSRRB) to start recruitment at 13 of the proposed study sites and have since received approval to recruit at a further 2 sites. We have established a network of research nurses across the country and recruitment is now progressing well at 15 different sites. Recruitment of cases for the study began in January 2003 and control recruitment began in May 2003. Full population-based recruitment of cases ceased for women diagnosed after 30 June 2005 but collection of biospecimens and limited epidemiological data will continue at key centres until June 2006. We have recruited a total of 1129 women with ovarian cancer (with an additional 35 women recruited since 1 July 2005 for the biospecimens extension) and 1056 control women. The recruitment, sample and data collection and processing systems have worked well and we are continuously monitoring our performance against our targets as outlined in the reports from Core Components A (Epidemiology) and B (Biospecimens)..

**15. SUBJECT TERMS**
Molecular epidemiology, heterogeneity, histologic subtypes, ovarian cancer

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON USAMRMC |
|---|---|---|---|---|---|
| a. REPORT U | b. ABSTRACT U | c. THIS PAGE U | UU | 18 | 19b. TELEPHONE NUMBER (include area code) |

# TABLE OF CONTENTS

# ANNUAL REPORT

| | |
|---|---|
| Title: | Molecular Epidemiology of Ovarian Cancer: The Australian Ovarian Cancer Study," |
| Proposal No. | OC000109 |
| Award No. | DAMD17-01-1-0729 |
| HSRRB Log No. | A-10884. |

## INTRODUCTION

The aim of this Program is to study the association between epidemiologic risk factors, low-risk genes, and histologic and novel molecular subtypes of ovarian cancer, thereby addressing the heterogeneity of the disease and of susceptibility to environmental exposures. To this end, we have established a multi-center population-based resource involving collection of linked epidemiologic and clinical data and biospecimens from cases and matched controls.

## BODY

### Cores A and B

In December 2002, we received final approval from the Human Subject Research Review Board (HSRRB) to start recruitment at 13 of the proposed study sites. During the final months of 2002 we identified Research Nurses for each study site and set up procedures at each hospital. By January 2003 research nurses were deployed at the collecting sites and we started phasing in recruitment. We now have HSRRB approval to recruit at all 15 sites listed below.

Recruitment and data/sample collection is proceeding well (see Reports for Cores A, Epidemiology, and B, Biospecimens, below).
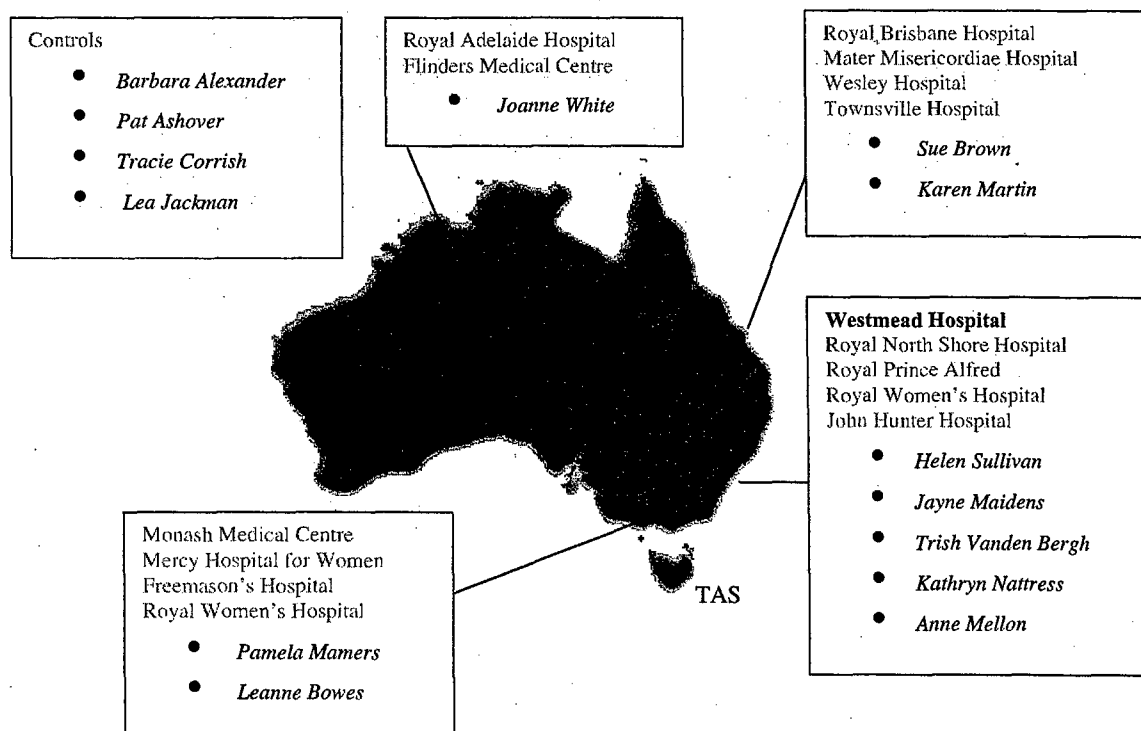


Figure 1: Australian Ovarian Cancer Study Site and RN network at 14 September 2005

<u>Tasks Outlined in the Approved Statement of Work</u>
<u>Core A: Epidemiology</u>

**Task 1- Preliminary Work (Prior to start date)**

(a)   Data-collection instruments will be finalised and piloted based on practical experience in a previous study (Survey of Women's Health)

Completed – see 2002 Annual Report.

(b)   Institutional Review Board approval will be obtained from all participating hospitals and institutions (to be prior to start date)

Completed – see 2003 Annual Report.

(c)   Identification of project manager, data manager and nurse-interviewers to start on Day 1.

Completed – see 2003 Annual Report.

**Task 2- Set-up (months 1-2)**

(a)   Finalise details of case identification system in each of the major centres (month 1)

Completed – see 2003 Annual Report.

(b)   Training of interviewers in Brisbane (month 1)

Completed – see 2003 Annual Report.

(c)   Development of computer data-bases (Access) for data-entry
      Recruitment "Tracking" Database

Completed – see 2003 Annual Report.

Main Questionnaire and Dietary Questionnaire Databases

Completed – see 2004 Annual Report.

**Task 3- Recruitment of cases (n>1000). (Ongoing months 2-36)**

(a)   Cases will be identified by the nurse-interviewers on an ongoing basis through participating hospitals and clinics with additional checks run through the state cancer registries
(b)   Treating physicians will be contacted to obtain permission to contact the case
(c)   Cases will be contacted and interviewed and biological samples collected
(d)   Tumor blocks and copies of pathology records will be obtained

Population-based recruitment was phased in from 1 January 2003 and continued to include women diagnosed up to June 30 2005. Although, for financial reasons, the recruitment period was 5 months shorter than initially planned, we have met our target of more than 1000 women and more than 600 fresh tissue samples (Table 1). Final

collection of outstanding questionnaires and biospecimens is ongoing and will be completed by late 2005.

NB. Recruitment and collection of biospecimens and limited epidemiological data will continue at most major centres until June 2006 under an extension to the original grant.

**Case ascertainment:** The total numbers of eligible cases in each state are estimated from numbers of ovarian cancers registered in each state in 2001 (the most recent data available). We were aiming to ascertain 100% of cases in Qld and SA and 50% in Victoria and New South Wales and have met these targets in all states except Victoria where recruitment at one major hospital did not start until 1 Jan 2004 (because of overlap with an existing study at that site).

**Response rate:** We were aiming for a response rate of 85-90% and, to date, 86% of cases who have been approached about the study (81% of those identified) have consented to take part.

**Data and sample collection:** Questionnaire return and blood, tissue and urine collection lag slightly behind consent thus these figures are under-estimates of final completion rates. Sample collection will be discussed further under the Biospecimen Core Report.

**Table 1: Case Recruitment: Cases diagnosed 1 Jan 2003 – 30 June 2005**

|  | NSW | QLD | SA | VIC | TOTAL |
|---|---|---|---|---|---|
| **Total cases (estimated)[1]** | 850 | 505 | 213 | 830 |  |
| **Cases ascertained** | 519 | 608 | 228 | 309 | 1664 |
| % of total (see Fig 2) | 61% | 120% | 107% | 37% |  |
| **Cases included[2]** | 431 | 488 | 191 | 280 | 1390 |
| % of those ascertained | 83% | 80% | 84% | 91% | 84% |
| **Consents[3]** | 356 | 378 | 146 | 249 | 1129 |
| % of included | 83% | 77% | 76% | 89% | 81% |
| **Q returned to QIMR[3]** | 295 | 326 | 112 | 182 | 915 |
| % of consents | 83% | 86% | 77% | 73% | 81% |
| **Fresh tissues[3]** | 207 | 166 | 81 | 158 | 612 |
| % of consents | % | 44% | 56% | 64% | 54% |
| **Blood samples[3]** | 176 | 350 | 103 | 227 | 956 |
| % of consents | 78% | 93% | 71% | 91% | 85% |
| **Urine samples[3]** | 206 | 147 | 0 | 95 | 448 |
| % of consents | 58% | 39% | 0% | 38% | 40% |

[1] Based on number of cases diagnosed in 2001 (most recent data available). We only aimed to ascertain 50% of cases in NSW and Vic

[2] Women were excluded if they were not able to give informed consent (eg non-English speaking, psychological reasons, too sick)

[3] Consent, questionnaire completion and sample collection can occur some weeks after surgery thus these numbers do not represent final figures.

**Task 4- Recruitment of population controls (n>1000). (Ongoing months 2-36)**

(a)     Potential controls will be selected at random through the Commonwealth electoral roll on a weekly basis and frequency matched by age and geographic region to the distribution of cases identified the previous week

(b)     Invitation letters will be sent to controls

(c)     Telephone follow-up of controls

(d)     Interview of controls and collection of blood and urine samples

We began control recruitment in late March 2003. To date 1056 women have consented to take part. [*Note: this represents 47% of those initially approached but does not allow for women who have been approached but have not yet responded and some ineligible women have to be excluded from the denominator. The final response rate will thus be considerably higher than 47%*]. We have therefore reached our target of 1000 control women. The questionnaire return rate is currently 92% and the blood collection rate is 78%. These are also not final figures because we do not yet have questionnaires/samples for women who only recently consented. The final return rates will thus be higher than this. Control recruitment will be completed by 30 September 2005.

**Control Recruitment 1 April 2003 – 14 September 2004**

|  | TOTAL |
|---|---|
| **Consents** | **1056** |
| **Q returned** | **973** |
| % of consents[1] | 92% |
| **Blood samples** | **825** |
| % of consents[1] | 78% |

[1] Questionnaires are returned and blood collected some time after consent is obtained thus these numbers do not represent final figures. [2] We are still setting up procedures for urine collection from control

**Task 5- Data entry / checking / cleaning. (ongoing months 3-42)**

(a)     Data will be entered into the databases on an ongoing basis

(b)     Data will be cleaned using frequency and range checks, implausible values will be cross-checked against the original questionnaires and corrected if necessary

Data entry is up to date and we have been conducting extensive data cleaning on an ongoing basis.

**Task 6- Data management (ongoing months 37-48)**

We are in the process of deriving key variables for analyses and anticipate having a complete and cleaned data-base for analysis by December 2005.

**Core B- Biospecimens**

**Task 1- Preliminary Work (Prior to start date)**

a)      **Institutional Review Board approval will be obtained from all participating hospitals and institutions (to be prior to start date)**

Completed – see 2003 Annual Report.

b)      **Recruitment of data manager and specimen processing staff**

Completed – see 2003 Annual Report.

c)      **Further refinement of computer data-bases (Access) for data-entry**

Completed- See 2002 Annual Report.

## Task 2- Set-up (months 1-2)

a)      **Finalise details of case ascertainment system in each of the major centres (month 1)**

Completed- See 2002 Annual Report.

b)      **Obtain minor equipment and consumables**

Completed- See 2002 Annual Report.

## Task 3- Ascertainment of samples. (Ongoing months 2-36)

a)      **Nurse-interviewers to liase with Biorepository head, notifying of incoming shipment of samples**

b)      **Nurse-interviewers to provide Biorepository staff with details of pathology blocks from cases to be requested.  Biorepository staff to coordinate temporary block acquisition**

c)      **Blood and urine samples shipped overnight from national sites.  Fresh frozen samples from interstate stored at centres at –80°C, shipped on dry ice at monthly intervals.  Blocks from pathology clinics requested on a monthly basis.**

As discussed above, a team of part-time nurse interviewers are in place at hospitals across Australia and sample collection for the biospecimen core has been in effect since January 1 2003.

Sample collection is going well. Blood samples are shipped at room temperature on a daily basis to the Biorepository and nurse-interviewers liase directly with the Biorepository head, notifying of incoming samples. This is usually via email. Fresh frozen tissue samples and urine samples are shipped on dry ice on a monthly basis via overnight courier. Again, the nurse-interviewers notify the Biorepository head of all incoming samples.

Nurse-interviewers collect both fresh frozen and fixed tissue samples. The fixed samples shipped to PMCC on a daily basis and are processed into blocks.

In addition, an expert gynaecological pathology review panel has been established to review all cases recruited to the study. Diagnostic slides and blocks from pathology centres (where the case was first diagnosed) are called in for centralized review. This protocol has been in place for several years to facilitate ovarian cancer research in Australia and is being co-ordinated by Dr Peter Russell and the Biorepository head.

Table 1 above shows the number of tissue, blood and urine samples collected since January 1 2003. As can be seen, we have met our initial target of more than 1000 women and more than 600 fresh tissue samples. Recruitment and collection of biospecimens and limited epidemiological data will continue at most major centres until June 2006 under an extension to the original grant. This additional funding will provide a unique opportunity to address some further questions that currently the study lacks sufficient power to investigate adequately.

## Task 4. Sample processing and dispatch. (Ongoing months 2-36)

a)      **Incoming samples of blood, urine, fresh frozen tissue and blocks to be processed as described in methods**
b)      **Requested samples shipped to centres in general on a monthly basis but immediately available if needed**
c)      **Sample backup to QIMR sent as batches on a monthly basis**
d)      **Periodic quality control procedures to validate sample integrity**

Joy Hendley and Lisa DiPrinzio are responsible for all biospecimen sample processing. Incoming blood, urine and tissue (frozen and fixed/blocks) samples are processed according to the AOCS protocols. A backup sample is stored separately (liquid nitrogen and -80°C) and sent via overnight courier to QIMR on a monthly basis.
We have established quality control protocols and the biorepository staff are responsible for their implementation.

## Task 5. Data entry / checking / cleaning. (ongoing months 3-42)

a)      **Data will be entered into the databases on an ongoing basis**
b)      **Provide data for analysis as required**

Biorepository staff are responsible for entering information regarding sample collection (type of sample, date collected, date processed) and processing (fractions processed, amount processed, storage location) onto the biospecimen database.

The Data Manager, Sian Fereday, is responsible for generating monthly statistics for the management group meetings. These reports describe primary site, histology subtype and stage of the biospecimens collected.

Adhoc data requests are furnished as required.

---

## Project 1: Molecular Subtype Analysis of Ovarian Cancer

## Task 1. Initial DNA microarray analysis with ~300 archival fresh frozen samples

**(months 1-12)**

Completed- See 2004 Annual Report

**Task 2. Progressively switch to microarray analysis of prospectively collected samples (months 12-42)**

The inclusion of prospective samples has commenced. The major focus of this project is an analysis of ~500 tumours using Affymetrix U133$^{2.0+}$ arrays. Analysis of the samples on spotted arrays provided valuable experience in developing SOP's for sample workup, including pathology review, extraction of RNA from frozen sections, RNA QC and requirements for clinical annotation before commencing sample analysis. An initial set of 360 mainly serous tumours has been selected, comprising approximately one-quarter archival samples and the remainder are AOCS prospectively collected samples. To date, 160 samples have been arrayed and a total of ~300 are expected to be completed by the end of 2005. This already represents the largest number of serous samples analysed on complex arrays to date.

Initial analysis of the expression data indicates that sample quality and array data is consistently excellent. Samples are frozen sectioned and reviewed for normal tissue, percent epithelial component, presence of necrotic tissue and consistency with the original pathology diagnosis. Additional sections are taken and are processed if the path review suggests that they are acceptable, that is they contain a high proportion of viable tumour with little normal tissue contamination (eg omentum or ovary). Samples are processed using Trizol and RNeasy and RNA quality accessed using a BioAnalyser, degradometer software, and a NanoDrop for quantitation. We follow quality control (QC) standards suggested by Affymetrix and by the Hammersmith Microarray Core facility, which has analysed over 2000 Affytmerix arrays. Almost all samples to date have had excellent to very good RNA profiles and all resulting Affymetrix arrays have passed QC measures.

The rate of sample processing depends on efficient coordination of multiple steps, including the accumulation of sufficient clinical information for sample inclusion, path review, sample processing and arraying. The major focus in the first 360 is relating expression patterns to overall survival and platinum resistance and these are expected to be completed by December 2005. Dr Anna Tinker is performing the current array analysis and we have hired a Research Assistant to work with her.

Preliminary analysis of Dr Tinker's data is currently in progress. Initial results suggest validation of the prediction of outcome data generated by Spentzos (JCO, 2004: 22, 4700-10). A subset of these samples (50) are being analysed by aCGH and the expression and copy number data will be combined (see below).

**Combined DNA copy number and gene expression analysis in serous ovarian cancer**

We are performing a pilot study of DNA copy number change (CNC) using 25 cases of Stage III serous ovarian cancer using Affymetrix SNP arrays, with a particular focus primary platinum resistance. Although we have previously performed cDNA aCGH, we

have switched the Affymetrix SNP platform due to their higher density representation of the genome and the ability to detect loss of heterozygosity (LOH) in regions where there is no net change in DNA copy number. Preliminary data has been generated on 15 tumours and 3 normal DNA reference controls using 50K XbaI SNP mapping arrays. Assay performance has been consistently high, as indicated by signal detection levels greater than 99% across all experiments. High SNP call rates were also observed in both the normal DNA reference and tumour samples. Consistently high call rates suggest that tumour sample purity has been sufficient to assign genotypes to a high proportion of SNPs and perform subsequent copy number analysis.

Initial analysis reveals a complex pattern of genomic change with many regions of amplification and LOH/loss in individual tumours. This includes amplified genes previously reported including PIK3CA, MYC and AKT2. Preliminary cross sample analysis has revealed regions of gain spanning the gene-rich region of 3q25.33-3q36.33. Further analysis on a larger set of samples with finer mapping of SNP arrays will assist in identifying the driving and co-amplified genes in large regions of gain such as at 3q. Amplification was also frequent along 8q23-24, including at SNPs flanking the MYC oncogene (67% of samples) although not exclusively spanning MYC, suggesting there may be other potential neighbouring oncogenes in this region. Contiguous point analysis (see Methods) has been applied to predict regions of LOH without reference to a matched normal sample. LOH events are frequent, including those on chromosome 17 which contain a number of tumour suppressor genes including p53, BRCA1, and OVCA1/2. LOH with and without copy number change can also be seen. This would not be detectable by BAC aCGH and presumably reflects regions of deletion followed by reduplication. The preliminary data from 15 samples demonstrates our ability to use the SNP arrays and to analyse data appropriately to identify copy number aberrations on a genomic scale.

We expect that as the sample size increases, common novel regions of CNC will be identifiable, as well as changes specific to the platinum resistant and responsive groups. Whilst our focus is on platinum resistance, we expect to obtain a wealth of other biological information about the molecular pathogenesis of ovarian cancer, which can be linked to other clinical parameters. For example, novel regions of amplification may be found that are not necessarily linked to platinum response but are involved in tumour aggressiveness or stage.

**Analysis of BRCA1 and BRCA2 tumours**
Approximately 10% of cases in population-based series carry germline mutations in BRCA1/2 and therefore that we can expect some of the AOCS cases, especially those with serous ovarian cancer, to be from mutation positive women. If there is an expression signature in tumours arising in such women, it may possible to identify women who might benefit from mutation testing, however, Jazaeri et al (JNCI 2002, 94: 990-1000) reported that mutation negative tumours resembled either BRCA1 or BRCA2 mutation positive tumours. We would like to recapitulate the Jazaeri et al study and have begun to accumulate tumours from carriers. To date, we have ovarian tumours from 15 BRCA1 carriers and 6 BRCA2 and are seeking further samples for analysis.

## Task 3. Ongoing statistical analysis of expression results (months 3-42)

Data will be hierarchically clustered to explore naturally appearing relationships, particularly when exploring samples with apparently similar clinico-pathological features such as an analysis of invasive serous cancers to search for unique molecular subsets.

Supervised data analysis can be used for other work where clinical or histological information separates groups, such as a comparison of LMP and invasive cancer or relating expression profiles to treatment response or survival. Currently there is no single 'best method' for associative analysis of genomic and clinical data and it is an area that is still under active development. We have a number of commercial software packages for microarray analysis and public domain tools that allow supervised analysis. In addition, we are working collaboratively with Dr Terry Speed's bioinformatics group, based at Walter and Eliza Hall Institute and with Dr Adam Kowalczyk's Statistical Machine Learning group, based at the National ICT Centre of Excellence (NICTA) in Canberra.

Genes will be tested using existing methods of feature-by feature statistical analysis including t-test, signal: noise metric, ANOVA and the score statistic (for Cox regression), with appropriate corrections for multiple testing using methods, in particular q-values, as well as empirical machine learning approaches. These methods are suited to an analysis of response that is considered as a categorical variable (resistant, sensitive). We will also treat outcome as a continuous variable, where disease free survival is used directly in the estimation process. For this analysis will use both standard tools (regression analysis, decision trees, neural networks, naïve Bayes), as well as a number of novel techniques developed in the Statistical Machine Learning Program at NICTA. The latter category includes advanced kernel methods and Bayesian techniques such as Gaussian Processes, Support Vector Machines, exponential families and graphical models. These models can incorporate prior knowledge, e.g. from the previous stages of the analysis. Predictive signatures of resistance will be tested using Kaplan-Meier analysis.

In analysing the data we will explore whether the analysis should be done with cases that represent clinical extremes, such as progression on primary treatment versus long-term survival, or whether the dataset can be analysed unselected. The former has the advantage that by choosing biological extremes, key genes may be more readily apparent. The disadvantage of this approach is that it relies on analysis of a smaller subset of samples than if the data was analysed without prior selection (beyond the initial classification of resistant and sensitive samples). In all analyses samples will be withheld by splitting the microarray data into multiple permutations of training and independent sets, and using leave one out cross validation. These methods should identify genes that are significantly correlated with outcome. Over expressed genes and regions of amplification or loss will be investigated using human reference genome assemblies such as Golden Path Genome Browser (University of California Santa Cruz) or Ensembl (EMBL - EBI and the Sanger Institute) to identify candidate oncogenes and tumour suppressor genes. To assess possible involvement of selected genes in related processes we determine whether specific GO annotations are enriched (http://www.geneontology. org/; http://david.niaid.nih.gov/david/ ease.htm) and use other commercial and public domain approaches to text and data mining to associate genes with pathways, such as Pathway Assist (http://www.stratagene.com/) and Ingenuity Pathway Analysis (http://www.ingenuity.com/). Techniques such as Gene Recommender will be used to identify genes that are correlated with members of the predictive list. Data will be

compared with large meta-analyses of different solid tumours to determine whether there are relationships between found genes and other data sets. We will test whether genes that are predictive of treatment response are correlated with expression patterns of response and survival for other tumours, such as lung (NSCLC), bladder, oesophageal cancer, where platinum is a mainstay of treatment. Where possible, we will determine whether our gene lists are associated with histological, clinical or biological parameters in other ovarian cancer datasets.

**Task 4. Full statistical analysis of expression data and preparation of manuscripts (months 42-48)**

See above

## Project 2: Determinants of Epithelial Ovarian Cancer- by histologic subtype and tumor behaviour

This project will commence formally when epidemiologic data collection and cleaning is complete at the end of the year however we have started preliminary work (see below) to ensure that we achieve our goals in a timely fashion. Analysis of the specific hypotheses will proceed in parallel under the guidance of the PI and Co-investigators.

**Task 1: Final preparation of data (month 1):**
(a)     **Final checking of dataset to eliminate outliers**
(b)     **Categorization of continuous variables (eg years of OC use) and generation of key derived variables (eg number of ovulations, body-mass index) to create standard variables for analyses**

Ongoing.

**Task 2: Testing of hypotheses relating to reproductive factors:**
(a)     **Analysis by histologic subtype**
(b)     **Analysis by tumor behavior (low malignant potential vs invasive)**
(c)     **Combination of (a) and (b) as required depending on previous results**
(d)     **Combining of new data with existing data from SWH**

Preliminary analysis has commenced using an interim dataset.

**Task 3: Testing of hypotheses relating to exogenous factors** (cigarette smoking, alcohol etc):
(a)     **Analysis by histologic subtype**
(b)     **Analysis by tumor behavior (low malignant potential vs invasive)**
(c)     **Combination of (a) and (b) if required**
(a)     **Combining of new data with existing data from SWH**

Preliminary analysis has commenced using an interim dataset.

**Task 4: Testing of hypotheses relating to health-related factors** (endometriosis, obesity etc):
(b)     **Analysis by histologic subtype**
(c)     **Analysis by tumor behavior (low malignant potential vs invasive)**
(d)     **Combination of (a) and (b) if required**

**(e)     Combining of new data with existing data from SWH**

Preliminary analysis has commenced using an interim dataset.

**Task 5: Testing of hypotheses relating to androgen exposure** (polycystic ovary syndrome, hirsutism, acne etc):
**(a)     Analysis by histologic subtype**
**(b)     Analysis by tumor behavior (low malignant potential vs invasive)**
**(c)     Combination of (a) and (b) if required**
**(d)     Combining of new data with existing data from SWH**

Preliminary analysis has commenced using an interim dataset.

**Task 6: Preparation of manuscripts for publication**

---

## Project 3: Low-risk genes for epithelial ovarian cancer

**Task 1. To establish the 16 single nucleotide polymorphism (SNP) genotyping assays, including identification of genotyping controls (months 1-18)**

**Task 2   To genotype the cases from the Survey of Women's Health Study and controls from the Australian Breast Cancer Family Study for 16 SNPS (months 6-24)**

Genotyping has been completed for ~550 ovarian cancer cases and 300 healthy controls for 11 SNPs as indicated in the table below. Four SNPs have been excluded from genotyping because of their low frequency (0-0.5%) detected in a sample of 90-125 Australian controls [published frequencies: 4% for *HSD17B1*:A-27C (Peltoketo et al, 1994); 2.4% for *RAD50*:Arg884His (http://greengenes/llnl.gov/dpublic/ secure/reseq) and 2% for *RAD52*:Ser347Ter (Han et al, 2002)].

| Gene | Polymorphism | Status |
|---|---|---|
| | | |
| Androgen Receptor (*AR*) | CAG$_n$ | Spurdle et al., 2002 |
| Progesterone Receptor (*PR*) | C44T | Not Commenced |
| Progesterone Receptor (*PR*) | G331A | Berchuck et al., 2004 |
| Aromatase (*CYP19*) | C>T 3'UTR | Completed |
| 5alpha-reductase (*SRD5A2*) | Val89Leu | Completed |
| 17βhydroxysteroid dehydrogenase (*HSD17B1*) | A-27C | Excluded |
| 17βhydroxysteroid dehydrogenase (*HSD17B1*) | Ala238Val | Completed |
| 17βhydroxysteroid dehydrogenase (*HSD17B1*) | Ser313Gly | Completed |
| 17βhydroxysteroid dehydrogenase (*HSD17B4*) | Trp511Arg | Completed |
| *BRCA2* | Asn372His | Auranen et al., 2003 |
| X-ray cross complementation (*XRCC2*) | Arg188His | Webb et al., 2004 |
| X-ray cross complementation (*XRCC3*) | Thr241Met | Webb et al., 2004 |
| X-ray cross complementation (*XRCC3*) | CA$_n$ | Not Commenced |
| *RAD50* | Arg884His | Excluded |
| *RAD52* | Ser347Ter | Excluded |
| *RAD52* | Tyr418Ter | Keleman et al, 2004 |

Analysis of the **RAD52** Y415Ter, **XRCC2** R188H G>A and **XRCC3** T241M C>T polymorphisms revealed no difference in genotype distribution between cases and controls. There was no increased risk of cancer associated with heterozygous genotype of *RAD52* Y415Ter (OR 0.55; 95% CI 0.24-1.24) (Keleman et al., 2004); *XRCC2* GA/AA genotype (OR 0.77; 95% CI 0.51-1.14), or with the *XRCC3* CT or TT genotypes (OR 0.80; 95% CI 0.59-1.09 and OR 0.92; 95% CI 0.58-1.44, respectively) (Webb et al., 2004). There was also no indication that genotype frequency differed across ovarian cancer subgroups defined by tumour characteristics, including histology. P53 positive tumours seemed to be over-represented in carriers of the *RAD52* truncation polymorphism (100% of 7 carriers but only 64% of 142 non-carriers were p53 positive). Although the rarity of the variant genotype frequency provided little power to detect modest risks in cancer for the *RAD52, XRCC2* and *XRCC3* variants, the data suggest that none of these variants play a major role in predisposition to ovarian cancer risk at the population level.

In contrast, in collaboration with Drs Easton et al in Cambridge, we have found an association between the Asn372His genotype of **BRCA2** and ovarian cancer risk (Auranen et al., 2003). We genotyped a total sample of 1121 ovarian cancer cases and 2643 controls. There was no difference in genotype frequency between control groups from the two Australian and British studies (P=0.9). The *HH* genotype was associated with an increased risk of ovarian cancer in both studies, and the risk estimate for the pooled studies was 1.36 (95% CI 1.04-1.77, P=0.03). There was also a suggestion that this risk may be greater for ovarian cancers of the serous subtype for both studies, with an OR (95% CI) of 1.66 (1.17-2.54) for the two studies combined (P=0.005). The *BRCA2* 372 *HH* genotype appears to be associated with an increased risk of ovarian cancer of a similar magnitude to that reported for breast cancer.

There was no evidence for a relationship between the variant allele and ovarian cancer risk for **CYP19** or **HSD17B1** [age adjusted OR (95%CI): *CYP19* 3'UTR heterozygous CT genotype 0.86 (0.60-1.24), *CYP19* 3'UTR homozygous TT genotype 0.70 (0.47-1.07); V allele of *HSD17B1* A238V 1.38 (0.35-5.49); heterozygous genotype of *HSD17B1* S313G 1.22 (0.87-1.72), homozygous GG genotype of *HSD17B1* S313G 0.98 (0.65-1.47)]. These odds ratios remained largely unchanged when tumours of low malignant potential (LMP) were excluded from the analysis. The genotyping protocol used for **HSD17B1** was used in the paper by Dunning et al (2004), on which Ms Livia Keleman (DoD-funded research assistant) is an author.

The +331G/A SNP in the progesterone receptor gene, **PR**, was genotyped in a population-based, case-control study from North Carolina (438 cases and 504 controls) and in 535 cases and 298 controls from the Survey of Woman's Health (Berchuck et al., 2004). This SNP, which is in the *PR* promoter, alters the relative abundance of the A and B isoforms and has been associated with an increased risk of endometrial and breast cancer. There was a suggestion of a protective effect of the +331A allele (AA or GA) against ovarian cancer in the North Carolina study (OR = 0.72, 95% CI = 0.47–1.10). Examination of genotype frequencies by histologic type revealed that this was due to a decreased risk of endometrioid and clear cell cancers (OR = 0.30, 95% CI = 0.09-0.97). Similarly, in the Australian study there was a non-significant decrease in risk of ovarian cancer among those with the +331A allele (OR 0.83, 95% CI = 0.51–1.35) that was strongest in the endometrioid/clear cell group (OR 0.60, 95% CI = 0.24–1.44). In the

combined US/Australian data that included 174 endometrioid/clear cell cases (166 invasive, 8 borderline) the +331A allele was significantly associated with protection against this subset of ovarian cancers (OR 0.46, 95% CI = 0.23–0.92). Preliminary evidence of a protective effect of the +331A allele against endometriosis also was noted in control subjects (OR = 0.19, 95% CI = 0.03-1.38). These findings suggest that the +331G/A progesterone receptor promoter polymorphism may modify the molecular epidemiologic pathway that encompasses both the development of endometriosis and its subsequent transformation into endometrioid or clear cell ovarian cancer.

We found that there was a trend for increased ovarian cancer risk associated with the L allele of the *SRD5A2* V89L polymorphism (age adj OR=1.30; 95% CI=1.03-1.62; $p_{trend}$=0.03) which was also apparent among the invasive tumours only (OR=1.25; 95% CI=0.98-1.59; $p_{trend}$=0.08). A significant trend was observed for reduced ovarian cancer risk associated with the R allele of the *HSD17B4* W511R polymorphism (age adj OR=0.68; 95% CI=0.47-0.97; $p_{trend}$=0.04); the trend remained similar for the invasive tumours (age adj OR=0.75; 95% CI=0.51-1.09; $p_{trend}$=0.14). Genotype frequency differences across ovarian cancer subgroups defined by tumour characteristics (including histology) are being explored.

**Task 3 To genotype the cases and controls from the Australian Ovarian Cancer Study for 16 SNPs (months 25-42)**

DNA has been extracted from a total of ~1468 cases and controls. Of these, ~1376 samples have been quantitated and diluted before plating 384 well plates. A further 245 case samples and 317 control samples from a related project, the Australian Cancer Study (PI - Professor Adele Green) are currently being quantitated and plated into 384 well plates. Genotyping of the AOCS cases and controls has commenced with initial genotyping of three SNPs in the Progesterone Receptor gene. Genotyping of the SNPs in the other genes listed above for all samples is to commence shortly.

**Task 4 To perform genotyping for 2 short tandem repeat (STR) polymorphisms on both case-control studies (months 36-42)**

Not started because blood collection is ongoing (see above).

**Task 5 Statistical analysis of the genotyping results from the Survey of Women's Health Study and controls from the Australian Breast Cancer Family Study (months 24-36)**

Not started because genotyping is not completed.

**Task 6 Full statistical analysis of the genotyping results (months 40-48) and preparation of manuscripts**

## PUBLICATIONS

Auranen, A., Spurdle, A.B., Chen, X., Lipscombe, J., Purdie, D.M., Hopper, J.L., Green, A., Healey, C.S., Redman, K., Dunning, A. M., Pharoah, P. D., Easton, D., Ponder, B.A.J.,

**Chenevix-Trench, G.**, and Novik, K.L. BRCA2 Asn372His polymorphism and epithelial ovarian cancer risk. *International Journal of Cancer* 103:427-30 (2003)

Dunning AM, Dowsett M, Healey CS, Tee L, Luben RN, Folkerd E, Novik KL, Kelemen L, Ogata S, Pharoah PD, Easton DF, Day NE, Ponder BA. Polymorphisms associated with circulating sex hormone levels in postmenopausal women. *J Natl Cancer Inst. 96(12):936-45 (2004)*

Berchuck, A., Shildkraut, J., Wenham, R.M., Calingaert, B., Ali, S., Henriott, A., Halabi, S., Rodgriguez, G.C., **Gertig, D.**, Purdie, D,M., Keleman, L., Spurdle, A.B., Marks, J. and **Chenevix-Trench**, G. Progesterone receptor +331A polymorphism is associated with a reduced risk of endometrioid and clear cell ovarian cancers. *Cancer Epidemiology, Biomarkers and Prevention* 13:2141-7 (2004)

**Webb, P.M.**, Hopper, J.L., Newman, B., Chen, X., Keleman, L., McCredie, M.R.E., Southey, M. Giles, G.G., **Chenevix-Trench,** G. and Spurdle, A.B. Double strand break repair gene polymorphisms and risk of breast or ovarian cancer. *Cancer Epidemiology, Biomarkers and Prevention* 14:319-23 (2005)

Kelemen, L., Spurdle, A.B., Purdie, D.M., **Gertig, D**. and **Chenevix-Trench**, G. *RAD52* Y415X truncation polymorphism and epithelial ovarian cancer risk in Australian women. *Cancer Letters* 218:191-7 (2005)

## KEY RESEARCH ACCOMPLISHMENTS

### Cores A and B
We have met our targets for patient recruitment and data/sample collection (See Reports for Core A Epidemiology and Core B Biospecimens). In relation to this we also have established systems to manage data and samples from all of the different sites.

### Project 1
Gene expression markers of potential prognostic significance have been identified and these need to validated on additional datasets. We have identified a set of genes with robust expression differences between LMP and invasive tumours. We have also developed a classification tool for Krukenberg and Krukenberg-like tumours, which is likely to be of considerable value for determining the origin of atypical ovarian tumours, especially mucinous ovarian tumours.

### Project 2
This project will not formally commence until epidemiologic data collection is complete at the end of the year. Preliminary work has however commenced using an interim dataset.

### Project 3
The most recent novel finding from the samples from the Survey of Women's Health, in collaboration with Dr Andrew Berchuck, is that the +331A allele of the PR gene is significantly associated with protection against endometrioid ovarian cancers (OR 0.46, 95% CI = 0.23–0.92).

## REPORTABLE OUTCOMES
N/A

## CONCLUSIONS
### Cores A and B
Recruitment of cases for the study began in January 2003 and control recruitment began in May 2003. Full population-based recruitment of cases ceased for women diagnosed after 30 June 2005 but collection of biospecimens and limited epidemiological data will continue at key centres until June 2006. We have recruited a total of 1129 women with ovarian cancer (with an additional 35 women recruited since 1 July 2005 for the biospecimens extension) and 1056 control women. The recruitment, sample and data collection and processing systems have worked well and we are continuously monitoring our performance against our targets as outlined in the reports from Core Components A (Epidemiology) and B (Biospecimens).

### Project 1
Initial array analysis on archival samples has been in progress throughout 2004 leading to the discovery of several marker genes. We have now switched further microarray expression analysis of ovarian cancer samples to an Affymetrix U1332.0+ array system. The intention is to array 500 samples, including the previously tested archival samples and the, as yet untested, prospective samples on U1332.0+ arrays, thereby creating a database of 100's of well-defined ovarian samples. To date, 360 AOCS cases have been selected (based on amount of clinical information available) and randomised. A total of 160 cases have been arrayed and initial GeneSpring analysis has commenced. We expect to have the first series completed the end of 2005.

### Project 2
This project will not formally commence until epidemiologic data collection is complete at the end of the year. Preliminary work has however commenced using an interim dataset.

### Project 3
Of the eleven polymorphisms nominated for analysis in this project, there is preliminary evidence from the SWH that four of them (in *PR, BRCA2, SRD5A2* and *HSD17B4*) are associated with ovarian cancer risk. Further analyses in the AOCS samples will provide independent testing of these SNPs in ovarian cancer risk, and if confirmed provide more power to look for associations with subtypes of ovarian cancer, and to start to look for gene-gene and gene-environment interactions

## REFERENCES
Spentzos D, Levine DA, Ramoni MF, Joseph M, Gu X, Boyd J, Libermann TA, Cannistra SA. Gene expression signature with independent prognostic significance in epithelial ovarian cancer. *J Clin Oncol. 2004 Dec 1;22 (23):4700-10.*

Jazaeri AA, Yee CJ, Sotiriou C, Brantley KR, Boyd J, Liu ET. Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. J Natl Cancer Inst. 2002 Jul 3;94(13):990-1000.

## APPENDICES
N/A

AD_____

Award Number: DAMD17-01-1-0729

TITLE: Molecular Epidemiology of Ovarian Cancer

PRINCIPAL INVESTIGATOR: David Bowtell, Ph.D.
Adele Green, Ph.D.
Georgia Chenevix-Trench, Ph.D.
Anna DeFazio, Ph.D.
Dorota Gertig, MBBS, Sc.D.
David Purdie, Ph.D.
Penelope Webb, Ph.D.
David Whiteman, Ph.D.

CONTRACTING ORGANIZATION: University of Melbourne
East Melbourne Victoria 3002

REPORT DATE: September 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20060315 076

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

| 1. REPORT DATE (DD-MM-YYYY) 01-09-2005 | 2. REPORT TYPE Annual | 3. DATES COVERED (From - To) 15 Aug 04 – 14 Aug 05 |
|---|---|---|

| 4. TITLE AND SUBTITLE Molecular Epidemiology of Ovarian Cancer | 5a. CONTRACT NUMBER |
|---|---|
| | 5b. GRANT NUMBER DAMD17-01-1-0729 |
| | 5c. PROGRAM ELEMENT NUMBER |

| 6. AUTHOR(S) David Bowtell, Ph.D., Adele Green, Ph.D., Georgia Chenevix-Trench, Ph.D., Anna DeFazio, Ph.D., Dorota Gertig, MBBS, Sc.D., David Purdie, Ph.D., Penelope Webb, Ph.D., David Whiteman, Ph.D. E-Mail: david.bowtell@petermac.org | 5d. PROJECT NUMBER |
|---|---|
| | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Melbourne East Melbourne Victoria 3002 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|

| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012 | 10. SPONSOR/MONITOR'S ACRONYM(S) |
|---|---|
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

The aim of this Program is to study the association between epidemiologic risk factors, low-risk genes, and histologic and novel molecular subtypes of ovarian cancer. In December 2002, we received final approval from the Human Subject Research Review Board (HSRRB) to start recruitment at 13 of the proposed study sites and have since received approval to recruit at a further 2 sites. We have established a network of research nurses across the country and recruitment is now progressing well at 15 different sites. Recruitment of cases for the study began in January 2003 and control recruitment began in May 2003. Full population-based recruitment of cases ceased for women diagnosed after 30 June 2005 but collection of biospecimens and limited epidemiological data will continue at key centres until June 2006. We have recruited a total of 1129 women with ovarian cancer (with an additional 35 women recruited since 1 July 2005 for the biospecimens extension) and 1056 control women. The recruitment, sample and data collection and processing systems have worked well and we are continuously monitoring our performance against our targets as outlined in the reports from Core Components A (Epidemiology) and B (Biospecimens)..

**15. SUBJECT TERMS**
Molecular epidemiology, heterogeneity, histologic subtypes, ovarian cancer

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON USAMRMC |
|---|---|---|---|---|---|
| a. REPORT U | b. ABSTRACT U | c. THIS PAGE U | UU | 18 | 19b. TELEPHONE NUMBER (include area code) |

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

# TABLE OF CONTENTS

# ANNUAL REPORT

| | |
|---|---|
| Title: | Molecular Epidemiology of Ovarian Cancer: The Australian Ovarian Cancer Study," |
| Proposal No. | OC000109 |
| Award No. | DAMD17-01-1-0729 |
| HSRRB Log No. | A-10884. |

## INTRODUCTION

The aim of this Program is to study the association between epidemiologic risk factors, low-risk genes, and histologic and novel molecular subtypes of ovarian cancer, thereby addressing the heterogeneity of the disease and of susceptibility to environmental exposures. To this end, we have established a multi-center population-based resource involving collection of linked epidemiologic and clinical data and biospecimens from cases and matched controls.

## BODY

### Cores A and B

In December 2002, we received final approval from the Human Subject Research Review Board (HSRRB) to start recruitment at 13 of the proposed study sites. During the final months of 2002 we identified Research Nurses for each study site and set up procedures at each hospital. By January 2003 research nurses were deployed at the collecting sites and we started phasing in recruitment. We now have HSRRB approval to recruit at all 15 sites listed below.

Recruitment and data/sample collection is proceeding well (see Reports for Cores A, Epidemiology, and B, Biospecimens, below).
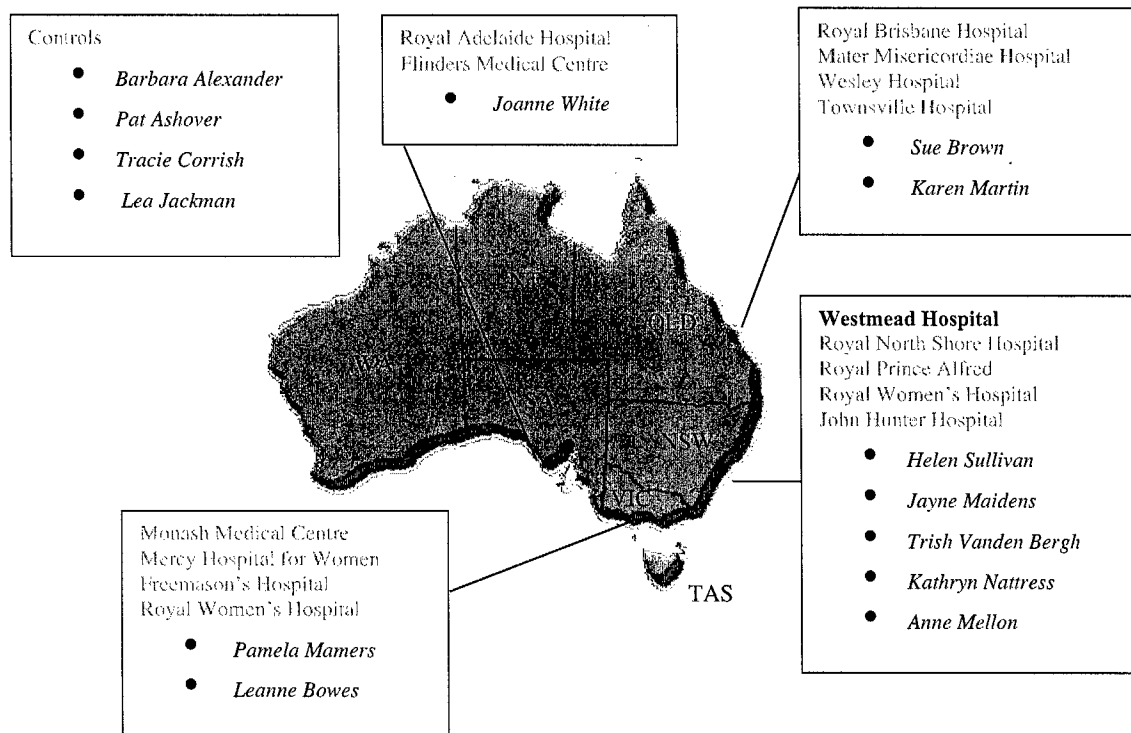


Figure 1: Australian Ovarian Cancer Study Site and RN network at 14 September 2005